

HW2: Neural Networks

1 Reading Exercises

From the `nnet.pdf` chapter, answer questions: 4.1, 4.2. CS6350, also answer 4.9.

2 Written Exercises

Answer the following questions in 25-100 words each:

1. Consider the perceptron update rule (from section 4.4.2 of the `nnet.pdf` reading) for binary classification (so the loss is 0/1 loss). What effect does the parameter η have on the learning process?
2. Suppose that we have N data points in D dimensions, *not* necessarily linearly separable. We can map this data into N points in $D + N$ dimensions, forcing it to be linearly separable. We may do this by mapping a the n th data point $\mathbf{x} = \langle x_1, x_2, \dots, x_D \rangle$ to $\langle x_1, x_2, \dots, x_D, [n = 1], [n = 2], [n = 3], \dots, [n = N - 1], [n = N] \rangle$, where by $[n = 3]$ we mean a value of 1 if $n = 3$ and a value of 0 otherwise. Show that the result of this mapping *is* linearly separable.
3. *[6350 only]* We proved in class (and in the write-up) that the perceptron algorithm will converge in R^2/γ^2 iterations, for linearly separable data with maximum norm R . Suppose we have linearly inseparable data, but we perform the mapping from the previous question. How long will the perceptron take to converge?